SPRAWOZDANIE

Zajęcia: Matematyka Konkretna

Prowadzący: prof. dr hab. Vasyl Martsenyuk

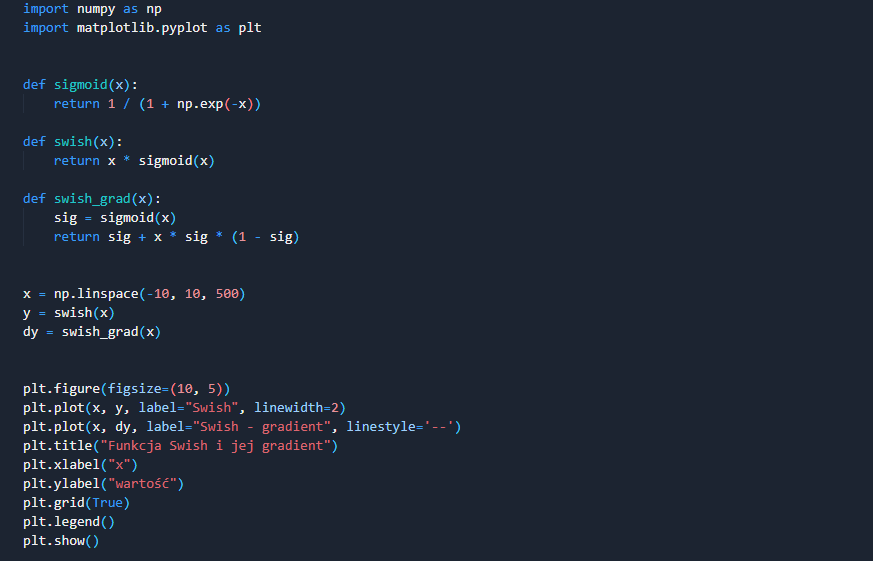
|  |  |
| --- | --- |
| Laboratorium Nr 4  Data 12.04.2025  Temat: „Funkcje aktywacji w sieciach neuronowych”  Wariant 11 | Anna Więzik  Informatyka  II stopień, niestacjonarne,  2 semestr, gr.1a TTO |

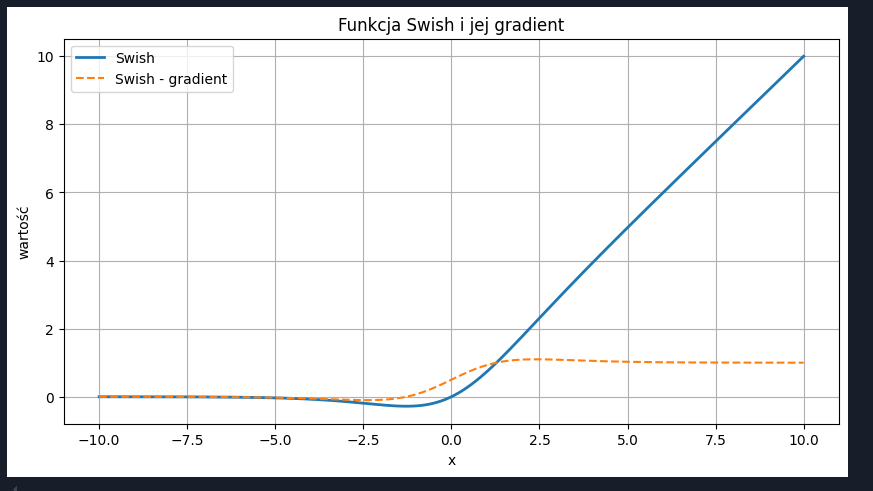
1. Polecenie:

Link do repozytorium: <https://github.com/AnaShiro/MK_2025>

Celem zadania było zbudowanie modeli uczenia maszynowego przy użyciu biblioteki Spark MLlib w celu przewidywania odpływu klientów banku na podstawie rzeczywistych danych. W ramach realizacji celu przeprowadzono pełny proces analityczny, począwszy od wczytania i przygotowania danych, przez eksplorację oraz ekstrakcję cech, aż po budowę i ocenę różnych modeli, takich jak regresja liniowa, K-means, drzewa decyzyjne i lasy losowe. Zastosowano również pipeline’y do automatyzacji przetwarzania danych i trenowania modeli oraz przeprowadzono strojenie hiperparametrów z użyciem walidacji krzyżowej.

1. Opis programu opracowanego





1. Wnioski

Na podstawie przeprowadzonych działań można wyciągnąć kilka wniosków. Przede wszystkim Spark MLlib okazał się skutecznym narzędziem do pracy z dużymi zbiorami danych i umożliwił sprawne tworzenie oraz testowanie modeli uczenia maszynowego. Użycie pipeline'ów znacząco uprościło proces przygotowania danych i trenowania modeli, a walidacja krzyżowa oraz tuning parametrów poprawiły jakość wyników. Modele klasyfikacyjne, takie jak drzewa decyzyjne i lasy losowe, umożliwiły przewidywanie, którzy klienci mogą opuścić bank, co może wspierać działania biznesowe związane z utrzymaniem klientów.